Disrupting malicious uses of AI by state-affiliated threat actors

Published: 2024-03-13 · Archived: 2026-04-05 17:19:49 UTC

Based on collaboration and information sharing with Microsoft, we disrupted five state-affiliated malicious actors: two China-affiliated threat actors known as Charcoal Typhoon and Salmon Typhoon; the Iran-affiliated threat actor known as Crimson Sandstorm; the North Korea-affiliated actor known as Emerald Sleet; and the Russia-affiliated actor known as Forest Blizzard. The identified OpenAI accounts associated with these actors were terminated.

These actors generally sought to use OpenAI services for querying open-source information, translating, finding coding errors, and running basic coding tasks.

Specifically:

- Charcoal Typhoon used our services to research various companies and cybersecurity tools, debug code and generate scripts, and create content likely for use in phishing campaigns.
- Salmon Typhoon used our services to translate technical papers, retrieve publicly available information on multiple intelligence agencies and regional threat actors, assist with coding, and research common ways processes could be hidden on a system.
- Crimson Sandstorm used our services for scripting support related to app and web development, generating content likely for spear-phishing campaigns, and researching common ways malware could evade detection.
- Emerald Sleet used our services to identify experts and organizations focused on defense issues in the Asia-Pacific region, understand publicly available vulnerabilities, help with basic scripting tasks, and draft content that could be used in phishing campaigns.
- Forest Blizzard used our services primarily for open-source research into satellite communication protocols and radar imaging technology, as well as for support with scripting tasks.

Additional technical details on the nature of the threat actors and their activities can be found in the Microsoft blog post(opens in a new window) published today.

The activities of these actors are consistent with previous red team assessments(opens in a new window) we conducted in partnership with external cybersecurity experts, which found that GPT-4 offers only limited, incremental capabilities for malicious cybersecurity tasks beyond what is already achievable with publicly available, non-AI powered tools_._
